Анализ эффективности алгоритмов обучения нечёткой сети Ванга-Менделя

О.П. Солдатова
Самарский национальный исследовательский университет им. академика С.П. Королева
Самара, Россия
soldatova.op@ssau.ru

И.А. Лёзин
Самарский национальный исследовательский университет им. академика С.П. Королева
Самара, Россия
lezin.ia@ssau.ru

И.В. Лёзина
Самарский национальный исследовательский университет им. академика С.П. Королева
Самара, Россия
lezina.iv@ssau.ru

Е.В. Муравьева
Самарский национальный исследовательский университет им. академика С.П. Королева
Самара, Россия
muraveva.ev@ssau.ru

***Аннотация* — При обучении нечётких сетей часто сталкиваются с проблемой выбора алгоритма обучения. Для решения проблемы выбора наиболее эффективного алгоритма обучения в работе были исследованы различные градиентные и стохастические алгоритмы обучения сети Ванга-Менделя. В качестве градиентных алгоритмов были выбраны алгоритм градиентного спуска и алгоритм сопряжённых градиентов. Стохастические алгоритмы представлены генетическим алгоритмом, алгоритмом имитации отжига, алгоритмом имитации роя частиц и алгоритмом дифференциальной эволюции. Эффективность алгоритмов обучения исследована на примере решения задачи классификации двух наборов модельных данных: ирисов Фишера и итальянских вин. Для исследования обучаются несколько модификаций сети Ванга-Менделя с различными алгебрами нечёткой логики: алгеброй Гёделя, алгеброй Гогена и алгеброй Лукашевича. Для классификации вин лучшие результаты показала модель сети с алгеброй Гогена, обученная генетическим алгоритмом. Для классификации ирисов с лучшие результаты также показала модель с алгеброй Гогена, обученная алгоритмом роя частиц и модель с алгеброй Лукашевича, обученная генетическим алгоритмом.**

Ключевые слова— нечёткая сеть Ванга-Менделя, нечёткая алгебра, алгоритм градиентного спуска, метод обратного распространения ошибки, алгоритм сопряжённых градиентов, генетический алгоритм, алгоритм имитации отжига, алгоритм имитации роя частиц, алгоритм дифференциальной эволюции.

I. Введение

Одной из задач, которую решают с помощью нечёткой логики является задача классификации объектов при пересекающихся классах. Для этих целей чаще всего используют нечёткие сети [1,2]. Во время обучения сети разработчик чаще всего сталкивается со следующими проблемами:

- недифференцируемость операций нечёткой логики, что ограничивает использование градиентных алгоритмов обучения;
- отсутствие информации о влиянии выбора алгебры нечёткой логики на эффективность решения задачи классификации.

Для решения первой проблемы применяют два способа:

- используют различные методы псевдо дифференцирования, например, для операции нечёткой конъюнкции в виде минимума, производная при обратном распространении ошибки считается только для связи с минимальным значением, остальные производные приравниваются нулю;
- используют стохастические алгоритмы обучения, не требующие дифференцируемости функций активации.

Для решения проблемы выбора нечёткой алгебры проводятся вычислительные эксперименты на моделях нечётких сетей [1,2]. В данной работе были проведены вычислительные эксперименты на моделях сети Ванга-Менделя с тремя различными алгебрами: Гёделя, Гогена и Лукашевича. Несмотря на множество методов и алгоритмов для решения данных проблем, справиться с ними в полной мере не удаётся. Поэтому исследования в этой области актуальны, как в настоящее время, так и будут актуальны в ближайшем будущем.

II. Нечёткий вывод и операции нечёткой логики

Сеть Ванга-Менделя реализует модель нечёткого вывода Мамдани-Заде, в которой присутствуют следующие операции [3]:

- операция нечёткой конъюнкции в виде логического или арифметического произведения для определения значения функции принадлежности условий правила;
- операция нечёткой импликации в виде логического или арифметического произведения для определения значения функции принадлежности заключения правила;
- операция нечёткой дизъюнкции в виде логической суммы для агрегации результатов импликации многих правил;
- оператор дефаззификации, трансформирующий нечеткий результат в четкое значение.

Таким образом, в модели Мамдани-Заде можно использовать правила вида:

$$\text{ЕСЛИ } x_1 \text{ это } A_1 \text{ И } x_2 \text{ это } A_2 \text{ И} \ldots \text{И } x_n \text{ это } A_n, \text{ ТО } y \text{ это } B \quad (1)$$

В работе исследуются модели сети Ванга-Менделя с отличными от классической модели Мамдани-Заде интерпретациями логических операций, заданными соответственно алгебрами Гёделя, Гогена и Лукашевича. Обозначим значения коэффициентов принадлежности условия и заключения правила вывода как $\mu_A(x)$ и $\mu_B(y)$ соответственно. Обозначим нечёткую конъюнкцию как «$\otimes$», нечёткую дизъюнкцию как «$\oplus$», а нечёткую импликацию как «$\rightarrow$». Тогда формулы (2) определяют операции в соответствии с алгеброй Гёделя, (3) – в соответствии с алгеброй Гогена, а формулы (4) – в соответствии с алгеброй Лукашевича [1]:

$$\begin{cases} \mu_A(x) \otimes \mu_B(y) = \min\{\mu_A(x), \mu_B(y)\} \\ \mu_{A\to B}(x,y) = \begin{cases} 1, \mu_A(x) \le \mu_B(y) \\ \mu_B(y), \mu_A(x) > \mu_B(y) \end{cases} \\ \mu_A(x) \oplus \mu_B(y) = \max\{\mu_A(x), \mu_B(y)\} \end{cases} \tag{2}$$

$$\begin{cases} \mu_A(x) \otimes \mu_B(y) = \mu_A(x)\mu_B(y) \\ \mu_{A\to B}(x,y) = \begin{cases} 1, \mu_A(x) \le \mu_B(y) \\ \mu_B(y) / \mu_A(x), \mu_A(x) > \mu_B(y) \end{cases} \\ \mu_A(x) \oplus \mu_B(y) = \mu_A(x) + \mu_B(y) - \mu_A(x)\mu_B(y) \end{cases} \tag{3}$$

$$\begin{cases} \mu_A(x) \otimes \mu_B(y) = \max\{0, \mu_A(x) + \mu_B(y) - 1\} \\ \mu_{A\to B}(x,y) = \min\{1, 1 - \mu_A(x) + \mu_B(y)\} \\ \mu_A(x) \oplus \mu_B(y) = \min\{1, \mu_A(x) + \mu_B(y)\} \end{cases} \tag{4}$$

III. Исследование эффективности алгоритмов обучения нечёткой сети Ванга-Менделя

В работе проведены исследования эффективности двух градиентных алгоритма обучения: градиентного спуска с методом обратного распространения ошибки (ГС), и алгоритма сопряжённых градиентов (АСГ) [3]. Также были исследованы четыре стохастических алгоритма: генетический алгоритм (ГА), алгоритм имитации отжига (АИО) [4], алгоритм имитации роя частиц (АИРЧ) и алгоритм дифференциальной эволюции (АДЭ) [5,6].

Исследования проводились на двух модельных наборах данных: ирисах Фишера и набора данных о химическом составе итальянских вин. Оба набора содержат данные трёх классов объектов [7]. Набор данных ирисов Фишера является сбалансированным, то есть для каждого класса существует одинаковое число примеров. Из 150 примеров 90 использовались для обучения и 60 для тестирования. Набор данных о химическом составе итальянских вин не является сбалансированным. Из 178 примеров 142 использовались для обучения и 36 для тестирования.

В качестве метрик качества решения задачи были использованы среднеквадратическое отклонение (СКО) тестирования и абсолютная погрешность классификации, равная отношению количества неверно распознанных данных при тестировании к общему объёму тестовой выборки. Результаты исследований эффективности алгоритмов обучения в зависимости от числа нейронов и числа эпох обучения для решения задачи классификации вин приведены в таблице 1.

Таблица I. Результаты исследований эффективности алгоритмов обучения

Тип алгоритма и алгебра	**Параметры и лучшие результаты тестирования**					
	Число нейронов	***Число эпох обучения***	***СКО тестирования***	***Число нейронов***	***Число эпох обучения***	***Абсолютная погрешность тестирования***
ГС, Гёделя	39	5000	0,164	65	1000	0,014
ГС, Гогена	**39**	**5000**	**0,006**	**39**	**500**	**0,000**
ГС, Лукашевича	39	1000	0,194	39	1000	0,028
ГА, Гёделя	65	500	0,143	65	500	0,028
ГА, Гогена	**117**	**1000**	**0,025**	**39**	**5000**	**0,000**
ГА, Лукашевича	39	5000	0,122	65	500	0,028
АИО, Гёделя	39	5000	0,191	39	5000	0,049
АИО, Гогена	**39**	**5000**	**0,138**	**91**	**5000**	**0,028**
АИО, Лукашевича	39	5000	0,278	117	5000	0,063
АИРЧ, Гёделя	39	50	0,174	39	50	0,042
АИРЧ, Гогена	**39**	**50**	**0,098**	**39**	**50**	**0,007**
АИРЧ, Лукашевича	39	50	0,282	39	50	0,113
АДЭ, Гёделя	39	5000	0,403	39	500	0,190
АДЭ, Гогена	**39**	**5000**	**0,273**	**65**	**100**	**0,092**
АДЭ, Лукашевича	39	1000	0,271	91	1000	0,120
АСГ, Гёделя	39	10	0,410	65	5	0,239
АСГ, Гогена	39	1	0,379	39	1	0,317
АСГ, Лукашевича	39	5	0,465	117	50	0,169

Как видно из таблицы, лучшие результаты даёт обучение моделей сети с алгеброй Гогена. Генетический алгоритм и алгоритм градиентного спуска обеспечили минимальные погрешности тестирования. При решении задачи классификации ирисов Фишера лучший результат обеспечил алгоритм имитации роя частиц с алгеброй Гогена, а также генетический алгоритм в сочетании с алгеброй Лукашевича.

Литература

[1] Солдатова, О.П. Решение задачи классификации с использованием нейронных нечётких продукционных сетей на основе модели вывода Мамдани-Заде / О.П. Солдатова, И.А. Лёзин // Вестник Самарского государственного технического университета. Серия физико-математические науки. – 2014. – Т. 2, №35. – С. 136-148. DOI: 10.14498/vagtu1266.

[2] Kirsh, D. 3D crystal structure identification using fuzzy neural networks / D. Kirsh, O. Soldatova, I. Lyozina, A. Kupriyanov, I. Lyozin // Optical Memory & Neural Networks (Information Optics). – 2017. – Vol. 26(4). – P. 249-256.

[3] Осовский, С. Нейронные сети для обработки информации / С. Осовский, пер. с польского И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.

[4] Глушенко, С.А. Обучение нейро-нечеткой сети с помощью генетического алгоритма / С.А. Глушенко, А.И. Долженко // Кибернетика и программирование. – 2017. – Т. 5. – С. 79-88.

[5] Скобцов, Ю.А. Эволюционные вычисления: учебное пособие / Ю. А. Скобцов, Д. В. Сперанский // Нац. открытый ун-т (ИНТУИТ). – М.: Нац. открытый ун-т (ИНТУИТ), 2015. – 326 с.

[6] Мостовой, Я.А. Оптимальное планирование операций роя подвижных объектов в условиях неопределённости // Я.А. Мостовой, В.А. Бердников // Компьютерная оптика. – 2020. – Т. 44, № 3. – С. 466-475. DOI: 10.18287/2412-6179-CO-599.

[7] Репозиторий UCI Machine Learning Repository [Электронный ресурс]. – Режим доступа: http://archive.ics.uci.edu/ml (09.09.2023)